



U.S. DEPARTMENT OF
ENERGY

Office of
Science

DOE Office of Science User Statistics Update

NUFO Annual Meeting
June 15th, 2016

Mariam Elsayed

Program Analyst

Office of the Deputy Director for
Science Programs

mariam.elsayed@science.doe.gov

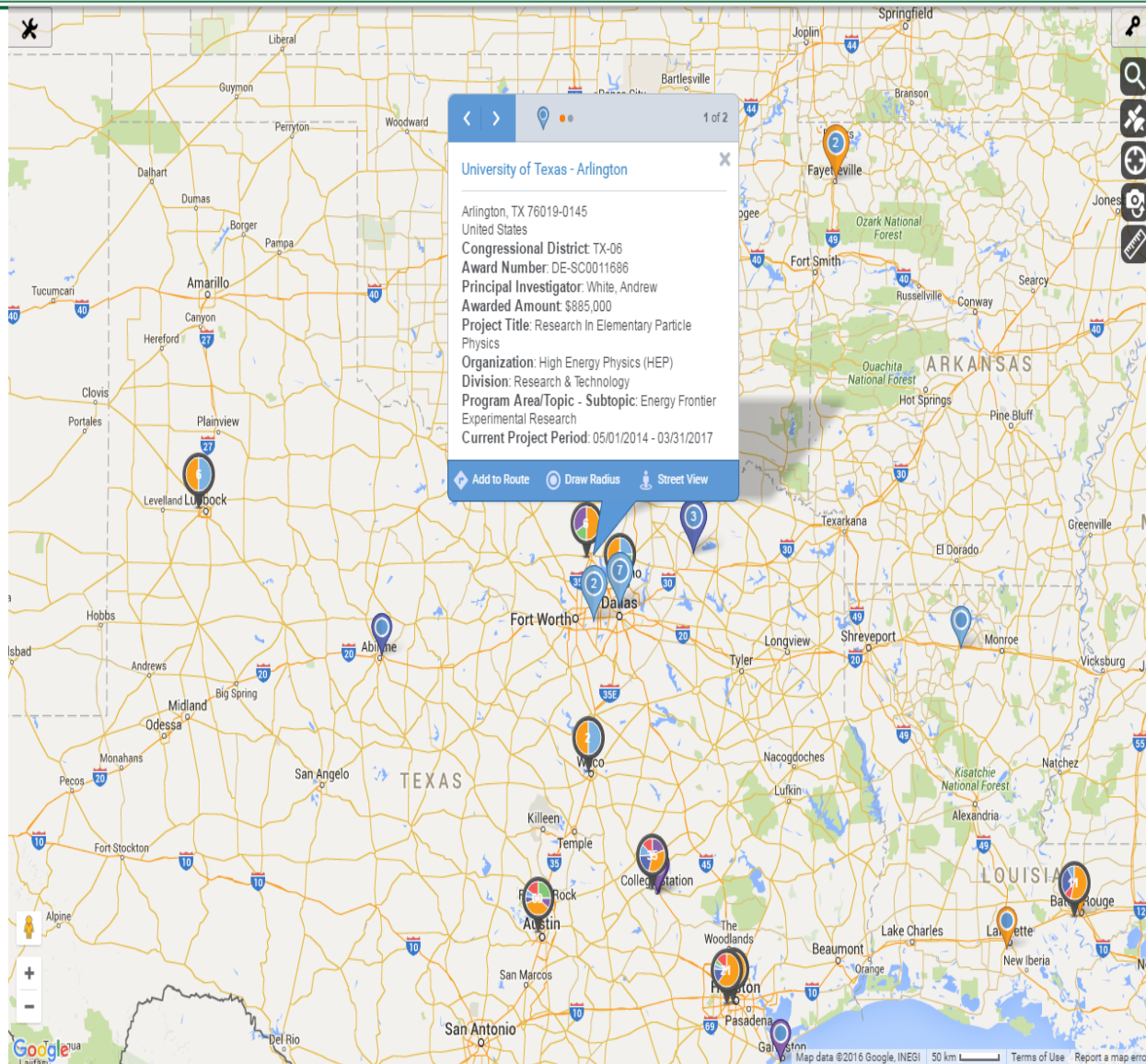
Outline

- **User Statistics**
 - Review of FY13-14 institution level statistics
 - FY2015 project level statistics (NOW LIVE!)
- **Lessons Learned**
 - Issues in Data Curation
 - Information Standards
 - Business Intelligence and Data Platforms
- **Outlook ahead and key takeaways**

FY2013 and FY2014 User Statistics

- **Initiative worked through the SC User Facilities Working Group and engagement with the SC User Facilities over two years.**
 - Defined user: Onsite/Remote/Data
 - Articulated user statistics collection practices
- **For these two years, user data was aggregated by institution.**

Interactive Grants Map



U.S. DEPARTMENT OF
ENERGY

Office of
Science

Feedback

- **Internal**
- **External**
 - Publications
 - Energy Sciences Coalition
 - National Lab Science Day on the Hill

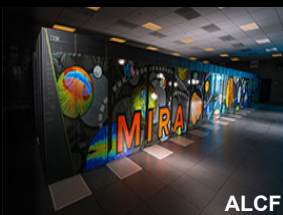


Starting from the left, SLAC director Chi Chang Kao, Representative Randy Hultgren (R-IL), Senator Mark Warner (D-VA), Senator James Risch (R-ID), Energy Secretary Ernest Moniz, Representative Bill Foster (D-IL) explore the interactive user statistics map demonstrated by Chris Smith (far right) at National Lab Science Day on the Hill. (Credit: Argonne National Laboratory)

FY 2015 28 user facilities



OLCF



ALCF



NERSC



ESnet



EMSL



ARM



JGI



SNS



HFIR



ALS



APS



LCLS



NSLS-II



SSRL



CFN



CINT



CNM



CNMS



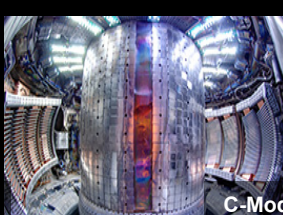
TMF



DIII-D



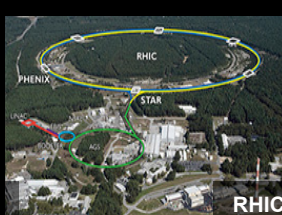
NSTX-U



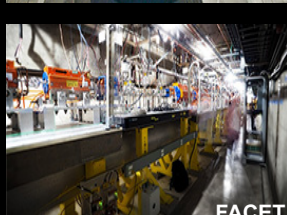
C-Mod



ATLAS



RHIC



FACET



ATF



Fermilab AC



CEBAF

User Statistics: Project Level

- **FY13/14 vs. FY15?**
 - Project Title
 - User Name
 - Funding Source
 - Employment Level (categories borrowed from BES survey)
 - User Type (On-Site vs. Remote)

FY2015 User Statistics

SC Home Organization Jobs Contact DOE Home

U.S. DEPARTMENT OF **ENERGY** Office of Science

Search SC Website SC Site Search GO

Programs Laboratories **User Facilities** Universities Funding Opportunities News About

You are here: SC Home » User Facilities » User Statistics

User Facilities

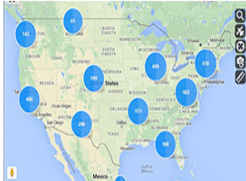
User Facilities Home
User Facilities at a Glance
User Resources
User Statistics
By Institution
By Project
Data Archive
User Statistics Collection Practices
Policies and Processes
Frequently Asked Questions
User Facility Science Highlights
User Facility News

User Statistics

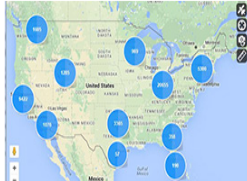
Print Text Size: A A Feedback (+) Share Page ▾

The Office of Science publishes annual statistics on the users of the Office of Science User Facilities, and provides interactive map representations of these user statistics. Prior year information is also available for download.

- User Statistics by Institution
- User Statistics by Project
- User Statistics Data Archive
- User Statistics Collection Practices



User Statistics by Institution Interactive Map



User Statistics by Project Interactive Map

Last modified: 5/26/2016 10:15:58 AM

CONTACT INFORMATION
Office of Science
U.S. Department of Energy
1000 Independence Ave., SW
Washington, DC 20585
P: (202) 586-5430

OFFICE OF SCIENCE
U.S. Department of Energy
1000 Independence Ave., SW

< > 1 of 2

University of Missouri - Columbia

United States
Program Full Name: Basic Energy Sciences
User Facility Full Name: Advanced Light Source
User Facility Host Institution Name: Lawrence Berkeley National Laboratory
Project/Experiment Title: Pyrogallol[4]arene: modified macrocycles for cocrystallization and manufacturing of self-assemblies
Project Type: Non-Proprietary
Funding Source: NSF
User Name: Feaster, Kyle
User Type: Remote
User Employment Level: Graduate Student

Add to Route Draw Radius Street View

<https://fortress.maptive.com/ver4/doescusersbyinstitutionfy2015>

<https://fortress.maptive.com/ver4/doescusersbyprojectfy2015>

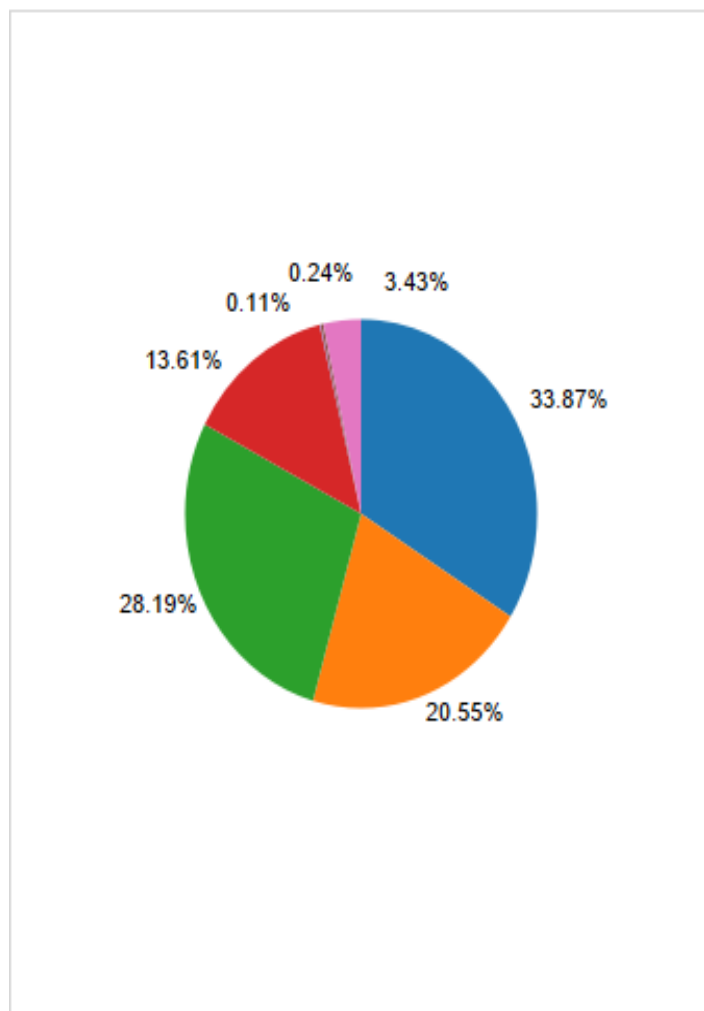


U.S. DEPARTMENT OF
ENERGY

Office of
Science

FY2015 User Statistics Analysis – Employment Level

Total



% of Total Total

100.00%

User Employment Level1

- Faculty member/Professional staff/Research scientist
- Graduate Student
- N/A
- Postdoctoral research associate
- Retired or Self Employed
- Student
- Undergraduate Student

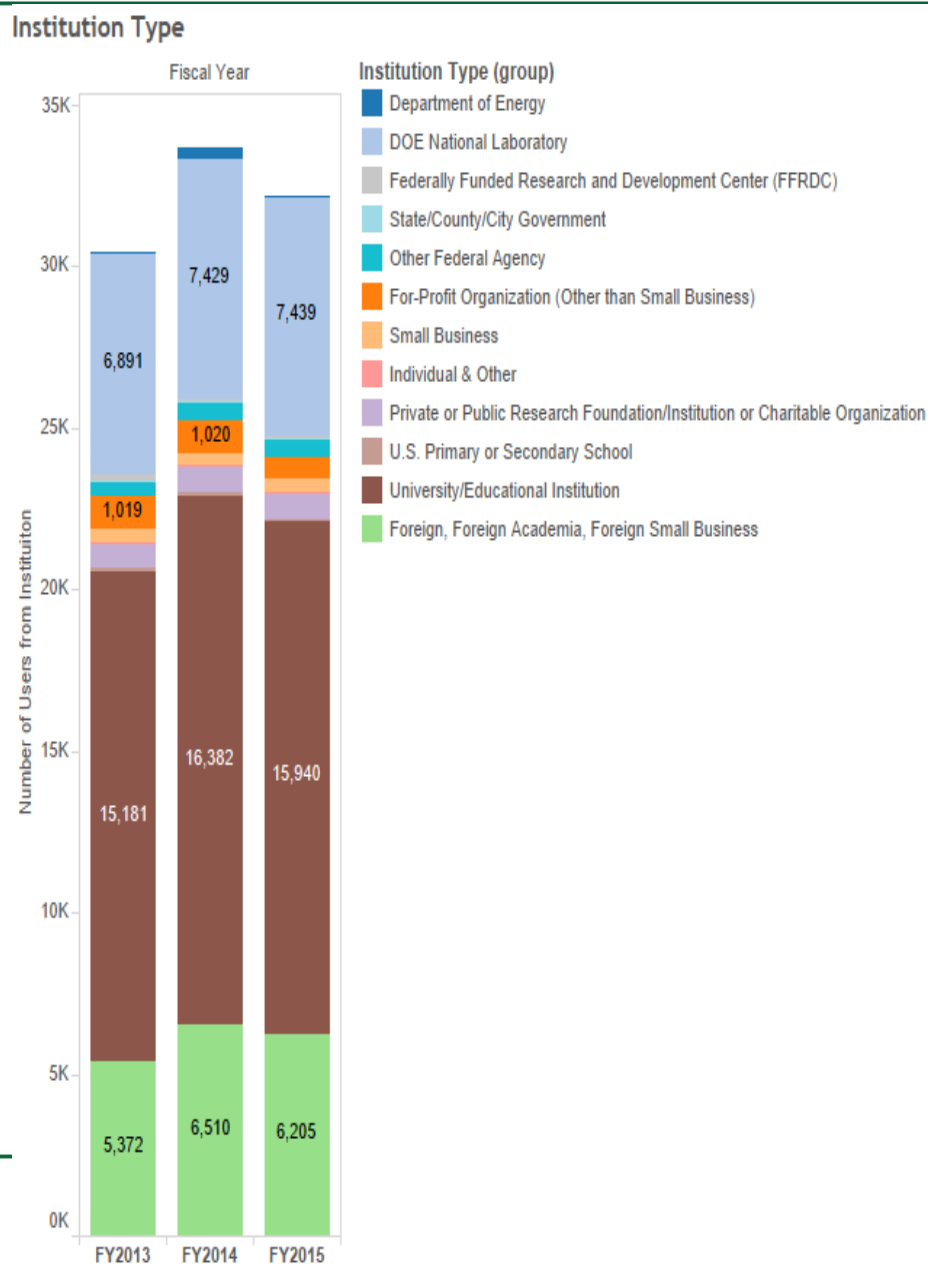
Note: These graphics are intended to preview some of the different analyses possible with the user statistics and are not final versions.



U.S. DEPARTMENT OF
ENERGY

Office of
Science

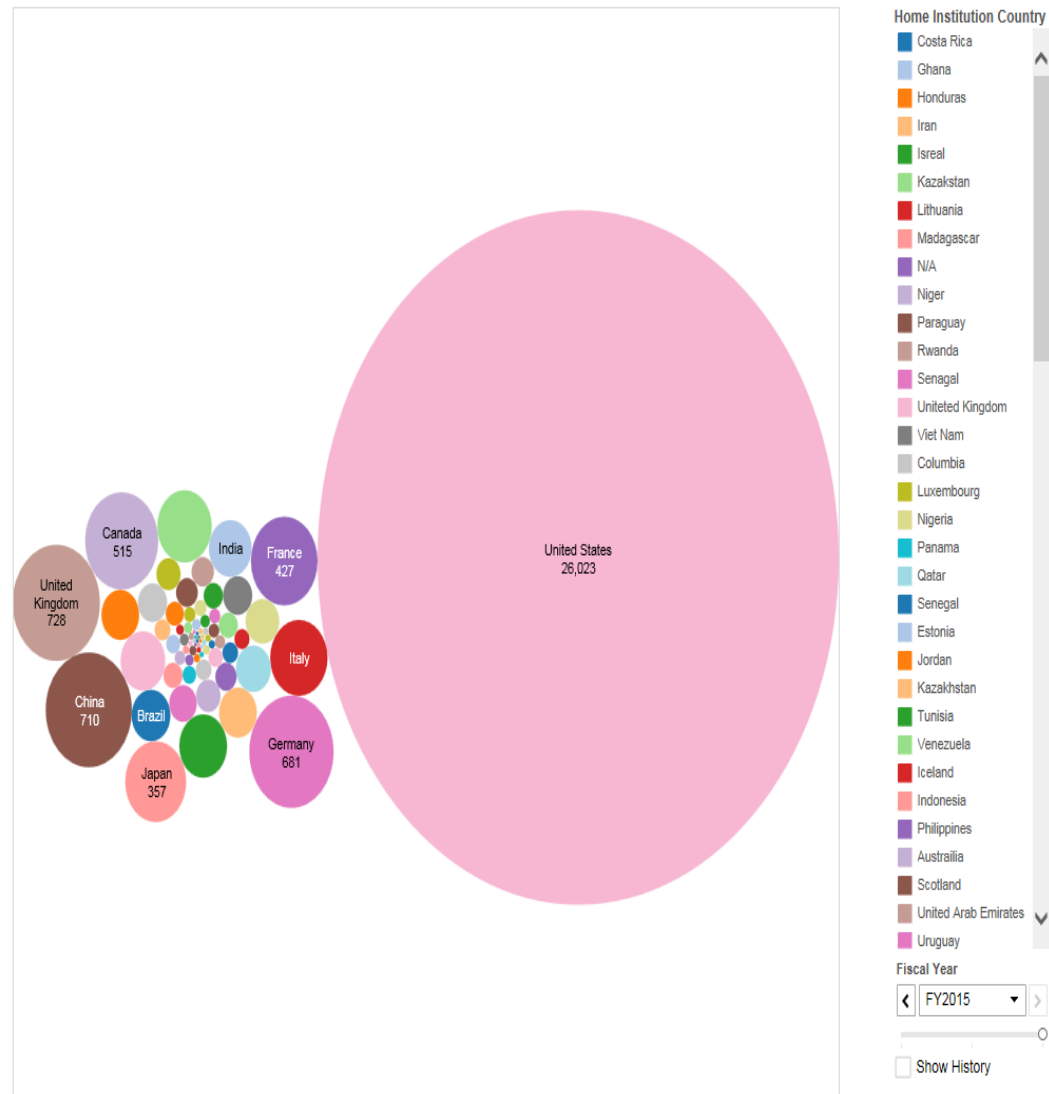
FY2015 User Statistics Analysis – Institution Types



Note: These graphics are intended to preview some of the different analyses possible with the user statistics and are not final versions.



FY2015 User Statistics Analysis – Institutions and Users by Countries



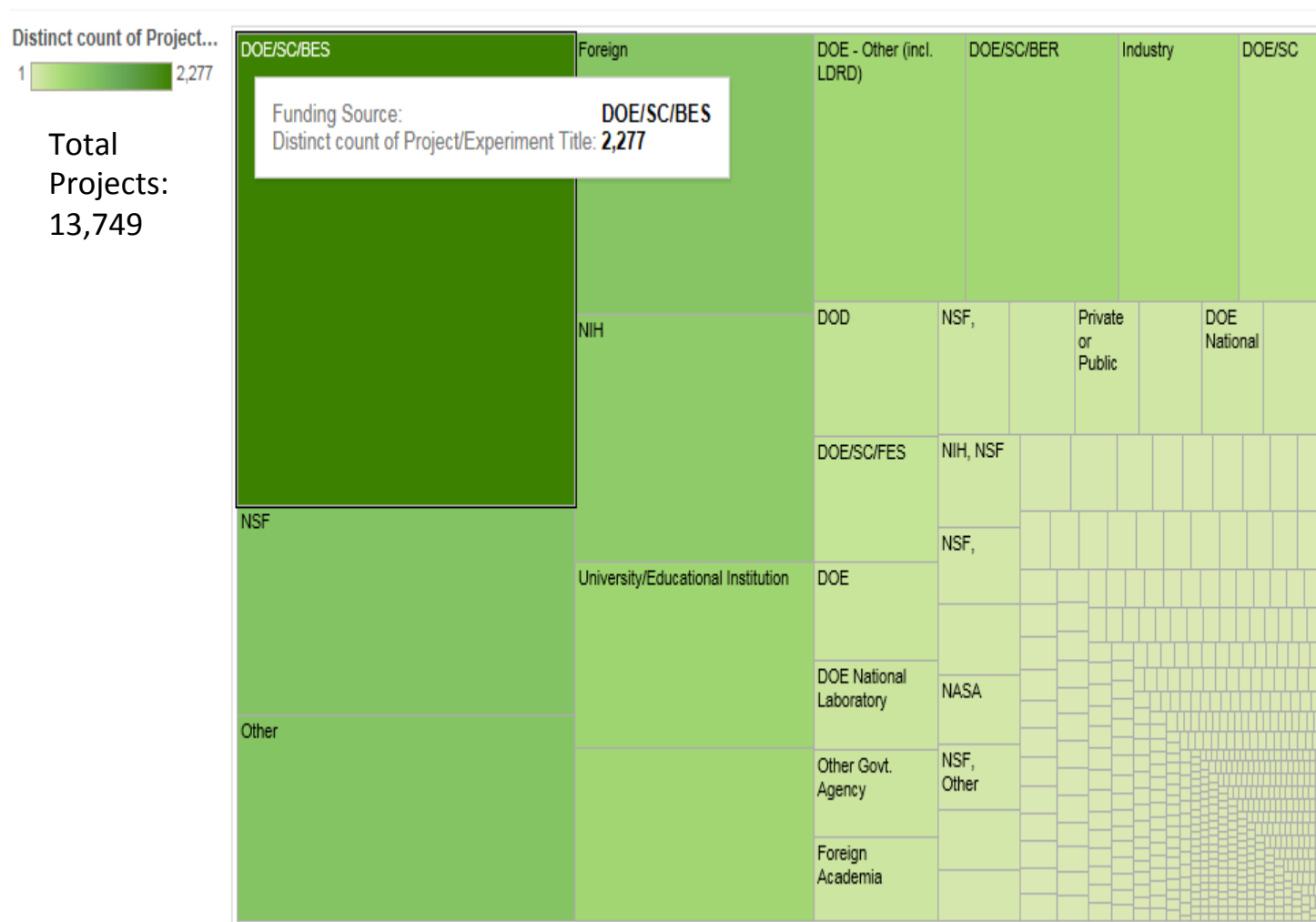
Note: These graphics are intended to preview some of the different analyses possible with the user statistics and are not final versions.



U.S. DEPARTMENT OF
ENERGY

Office of
Science

FY2015 User Statistics Analysis – Funding Sources



Note: These graphics are intended to preview some of the different analyses possible with the user statistics and are not final versions.



U.S. DEPARTMENT OF
ENERGY

Office of
Science

FY2015 User Statistics Analysis – User Crossover

FY 2015 Office of Science User Facilities -- user crossover analysis, June 2nd, 2016

Each element in the matrix represents the number of users that used the two facilities in that column and row. The data presented here should be interpreted as a lower bound as some facilities were not able to provide information that would have enhanced the accuracy of the analysis. Users were matched first according to email address; for those without email address, an attempt was made to resolve according to full name and institutional affiliation. Columns with a red box are the facilities that did not provide a complete record of user email addresses.

	Alcator C-Mod	ALCF	ALS	APS	ARM	ATF	ATLAS	CEBAF	CFN	CINT	CNM	CNMS	DIII-D	EMSL	Esnet	FACET	Fermilab AC	HFIR	JGI	LCLS	NERSC	NSLS-II	NSTX-U	OLCF	RHIC	SNS	SSRL	TMF	
Alcator C-Mod	224	2	1										48									40		35	2				
ALCF	2	990	4	18	9		1	5	2	3	15		3	12	1	1	10	3		4	281		2	169	2	4	3	2	
ALS	1	4	2560	286	13			6	6	9	8	4		18			2	13	2	65	90	5			1	15	223	96	
APS		18	286	5471			8	14	19	8	154	29		19		2	5	77	3	77	26	25	1	1		163	183	5	
ARM		9	13		1121			2						23			2				71			15			3		
ATF						75		1								10					4				2				
ATLAS		1		8			392	5		1	5										3			1					
CEBAF		5	6	14	2	1	5	1510	1	1	7	3		3			36	2		1	44		14	29	7	4	1		
CFN		2	6	19				1	493	1	2	3		1					3	1	19	12		2		3	3	2	
CINT		3	9	8			1	1	1	502		9						2		6	1		4		2	8	2		
CNM		15	8	154			5	7	2		529			2			4	2		4	16	2			1		8	1	
CNMS			4	29				3	3	9		575				1	1	29	1	16			3		74	1	2		
DIII-D	48	3											557					1			86		87	14				1	
EMSL		12	18	19	23			3	1		2			713				2	2	39	64			13		3	12	3	
Esnet		1													48						2								
FACET		1		2		10						1				148	1				3	10				1	3		
Fermilab AC		10	2	5	2			36			4	1	1	2			1924		2	1	45		3	4	81	3	2	2	
HFIR		3	13	77				2	3	2	2	29		2				492	2	2	4			3		235	7		
JGI			2	3								1		39				2	2	957	1	8				1	2		
LCLS		4	65	77				1	1		4					3	1	2	1	829	14	3		1		1	93		
NERSC	40	281	90	26	71	4	3	44	19	6	16	16	86	64	2	10	45	4	8	14	6332	2	41	304	142	13	14	30	
NSLS-II			5	25					12	1	2									3	2	95					1	6	1
NSTX-U	35	2		1									87					3			41		356	4					
OLCF	2	169		1	15		1	14	2	4		3	14	13				4	3		1	304		4	1107	1	4		
RHIC		2	1			2		29				1					81				142			1	1015	1			
SNS		4	15	163				7	3	2		74		3		1	3	235	1	1	13	1		4	1	843	7		
SSRL		3	223	183	3			4	3	8	8	1		12		3	2	7	2	93	14	6					1626	31	
TMF		2	96	5				1	2	2	1	2	1	3			2		2		30	1					31	677	
Total Crossover Use	128	556	867	1123	138	17	24	186	80	57	231	176	240	216	3	32	207	386	61	271	1395	58	173	559	260	538	613	179	
Percentage from to	57.14%	56.16%	33.87%	20.53%	12.31%	22.67%	6.12%	12.32%	16.23%	11.35%	43.67%	30.61%	43.09%	30.29%	6.25%	21.62%	10.76%	78.46%	6.37%	32.69%	22.03%	61.05%	48.60%	50.50%	25.62%	63.82%	37.70%	26.44%	

Total Users: 32161



U.S. DEPARTMENT OF
ENERGY

Office of
Science

Data Synthesis & Curation Challenges

- **User Identifiers**

- Email addresses are not always effective at resolving individual users

- **Standardization & Normalization**

- Institution Names
- Institution Addresses
- Institution Types
- Funding Sources



Information Standards

- **ORCID**
- **Funding Agency**
- **Institution profiles**
 - Names
 - Addresses
 - Geocoordinates
 - Institution types
 - Industrial categorization



Open letter from publishers

- <https://orcid.org/content/requiring-orcid-publication-workflows-open-letter>

Requiring ORCID in Publication Workflows: Open Letter

Major publishers have committed to requiring ORCID iDs in the publishing process for their journals and invite other publishers to do the same.

In November, 2015, a group of publishers asked ORCID to help facilitate communications about their plans to require authors to use an ORCID iD, including hosting this open letter explaining their rationale, developing best practices for using iDs in publishing, and maintaining the signatory list. The publishers' goal is to encourage others to join them in supporting the adoption of ORCID. Publishers signing this open letter are committing to requiring ORCID iDs during 2016 following specific implementation standards.

Why Require ORCID?

ORCID provides a digital name – or iD – that uniquely and persistently identifies researchers and contributors. By connecting this iD to different research activities and affiliations across multiple research information platforms, ORCID helps enable recognition and reduce reporting burdens for researchers. It also enhances the discovery process and lays the foundation for trust in a digital research environment. Crossref's recently launched Auto-Update functionality – which (with researcher permission) updates an author's ORCID record when the author uses their iD during the publishing process – further benefits researchers by alleviating the need to re-enter publication data in multiple systems. Importantly, it also provides an independent assertion of the connection between the author and their work.

With more than 3,000 journals already collecting ORCID iDs from corresponding authors, through all major manuscript submission systems, publishers are in a unique position to facilitate widespread adoption of ORCID. Indeed, researchers are most likely to encounter ORCID in publishing systems: about 75% of registrations occur because journals are asking authors to include their ORCID in new submissions. Researchers are also increasingly encountering ORCID in their university systems, and funders including the Wellcome Trust and EMBO have begun to require the use of ORCID in grant application systems.



U.S. DEPARTMENT OF
ENERGY

Office of
Science

ORCiD Members



U.S. DEPARTMENT OF
ENERGY

Office of
Science

ORCiD – Federal Use



- Optional field with Graduate Research Fellowship Program profiles
- Optional field in Fastlane



National Institutes
of Health

- SciENcv
- PubMed



- Optional field in E-Link, SciTech Connect



U.S. DEPARTMENT OF
ENERGY

Office of
Science

Funding Agencies

- **How to analyze with multiple funding sources?**



Institution Profiles

- **Names**
 - U.S. Dept. of Education Accredited Institution list of colleges and universities
 - Hoovers
 - Crossref
 - ORCID?
- **Location**
 - Addresses
 - Geocoordinates
- **Institution types**
 - PAMS
- **Industrial categorization**
 - OTT
 - NAICS codes

Business Intelligence and Data Platforms

- **Assessing various tools BI tools, cloud platforms, or separate architecture which can allow for data synthesis and analysis**
- **We are continuing to learn and do research and would appreciate your insights.**

Outlook

- **Considering building a data-ingestion portal/tool**
- **We think this tool could facilitate communication/interaction between key stakeholders who wrestle with user statistics**
- **Feedback and engagement is very important**
- **Information Standards**

**FY 2016
28 user facilities**



OLCF



ALCF



NERSC



ESnet



EMSL



ARM



JGI



SNS



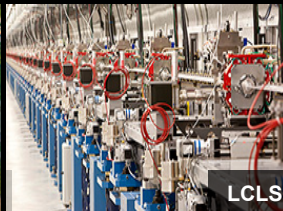
HFIR



ALS



APS



LCLS



NSLS-II



SSRL



CFN



CINT



CNM



CNMS



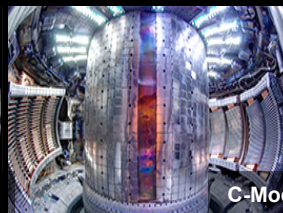
TMF



DIII-D



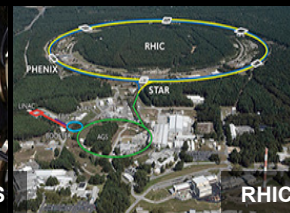
NSTX-U



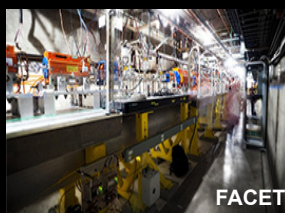
C-Mod



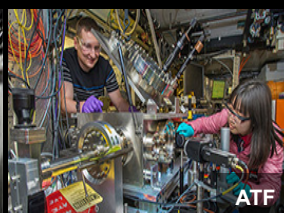
ATLAS



RHIC



FACET



ATF



Fermilab AC



CEBAF

Questions?

Mariam Elsayed

Mariam.Elsayed@science.doe.gov

(301) 903-0802